NIKHEF 96-02

Discrepancy-based error estimates for Quasi-Monte Carlo.

I: General formalism

Jiri Hoogland[1] **and Ronald Kleiss**[2]
NIKHEF-H, Amsterdam, The Netherlands

Abstract

We show how information on the uniformity properties of a point set employed in numerical multidimensional integration can be used to improve the error estimate over the usual Monte Carlo one. We introduce a new measure of (non-)uniformity for point sets, and derive explicit expressions for the various entities that enter in such an improved error estimate. The use of Feynman diagrams provides a transparent and straightforward way to compute this improved error estimate.

Preprint submitted to Computer Physics Communications

[1]e-mail: t96@nikhefh.nikhef.nl, research supported by the Stichting FOM
[2]e-mail: t30@nikhefh.nikhef.nl, research supported by the Stichting FOM

1 Introduction

In numerical integration, an object of paramount importance is the integration error, that is, the difference between the actual (unknown) value of the integral and the obtained numerical estimate. Since exact knowledge of the error would imply exact knowledge of the integral (and hence would obviate the need for a numerical estimate), in practice one has to settle for an estimate of this error, which must be shored up by plausibility arguments. In deterministic integration, where the set X_N of N integration points is determined beforehand, a deterministic error bound (that is, a true upper limit on the error) can often be obtained, provided the integrand falls in some known class of functions with, for instance, given smoothness properties. The integration rule (*i.e.* the point set X_N) can then be optimized to guarantee a rapid decrease of the error with increasing N [1].

Unfortunately, in many practical problems the integrand is not particularly smooth. This happens, for instance, in particle phenomenology, where integrands may contain discontinuities corresponding to experimental cuts in the admissible phase space. In such cases, one usually relies on *Monte Carlo* integration, where the point set X_N is assumed to be chosen from the ensemble of random point sets[1]. In the derivation of the error estimate, we use the fact that the point set X_N is a 'typical' member of the ensemble of all such point sets, and average over this ensemble. The error estimate is, then, probabilistic rather than deterministic; but with some care reliable estimates can be obtained, as implied by the various laws of large numbers. The Monte Carlo method can be applied to the very large class of square integrable integrands; on the other hand, there is the drawback that the error decreases asymptotically only as $1/\sqrt{N}$. The existing variance-reducing techniques can at most decrease the coefficient of the error estimate, and not its behaviour with N, which is a direct consequence of the random structure of the point set.

Recently, the approach known as *Quasi-Monte Carlo* has received considerable attention. Here, one attempts to construct points sets X_N that are 'more uniform' (in terms of a given measure of uniformity) than truly random ones. The existence of error bounds à la Koksma-Hlawka [2] underlies the belief that such *quasi-random* point sets can, indeed, lead to an error that decreases faster than $1/\sqrt{N}$: the Roth bound [3] suggests that a behaviour as $(\log N)^c/N$ may be possible, where c is some constant depend-

[1] In practice, the points are *pseudo-random* rather than truly random, being generated by a deterministic algorithm: however, we shall assume that the random number generator employed is sufficiently well behaved to satisfactorily mimic a truly random sequence.

ing on the dimensionality of the integral. Many quasi-random sequences with asymptotically small non-uniformity have been proposed, such as van der Corput/Halton [4], Faure [5], Sobol'[6], and Niederreiter [7] sequences. These generally perform to satisfaction in many applications: but, and this is the central problem addressed in this and subsequent papers, no reasonable way is yet known to estimate the error while making use of the improved uniformity of the point set. In fact, it is customary to just compute the error estimate *as if* the point set X_N were truly random. This procedure tends to overestimate the error, as shown by various case studies [8], so it is certainly safe to do so: but, in our opinion, *Quasi-Monte Carlo methods will come into their own only when improved error estimates are available.* It is our purpose to provide an attempt at such an improved integration error estimate. To do this, we shall have to face the fact that quasi-random point sets are *not* 'typical' members of an obvious ensemble: indeed, they are very special, with a lot of ingenuity going in their construction. The central idea of our approach is the definition of a suitable ensemble of quasi-random point sets.

The lay-out of this paper is as follows. In section 2, we shall discuss the error estimate for Monte Carlo, in such a way that it becomes clear precisely which information on the quasi-random point set is needed to improve the error estimate. To this end, we shall have to employ *some* definition of (non-)uniformity, that is, a *discrepancy.* In section 3, we shall introduce a particular definition of a discrepancy, which we feel is better suited to this kind of problem than the best-known discrepancy, the so-called *star discrepancy* D_N^*, discussed extensively in the literature [12]. In section 4, we shall discuss in detail how the various ingredients in a new error estimate can be computed, for a point set with a given, known, value of the discrepancy. We shall do this by essentially combinatorial arguments, using a technique based on Feynman diagrams. It will appear that, in the limit of large N, our approach is closely related to the computation of path integrals in field theory.

Finally, we have to apply in practice what we have learned in theory. We shall make such attempts, in the academic but instructive case of one-dimensional integral, and in more dimensions: these points will be treated elsewhere [9].

2 Estimates, ensembles and discrepancy

2.1 The improved error estimate

To set the stage, we define the D-dimensional integration region K to be $[0,1)^D$, the archetypical unit hypercube. The integrand is a function $f(x)$ defined on K, which may contain discontinuities but must be at least square integrable. Its integral is denoted by J_1, where

$$J_p = \int\limits_K dx \ (f(x))^p \quad , \quad p = 1, 2, \ldots \quad . \tag{1}$$

The point set X_N consists of N D-dimensional points x_k, $k = 1, 2, \ldots, N$. Where necessary, we shall denote individual components of points by Greek indices, x_k^μ, $\mu = 1, 2, \ldots, D$. The numerical estimate of the integral is then given by

$$S = \frac{1}{N} \sum_{k=1}^{N} f(x_k) \ , \tag{2}$$

and the error made is then simply $\eta \equiv S - J_1$. The salient fact about Monte Carlo is that the point set X_N is a random object, and so, therefore, is the error η. The standard Monte Carlo estimate is derived by assuming that X_N is a 'typical' member of the ensemble of point sets, governed by a probability density $P_N(x_1, x_2, \ldots, x_N)$. We shall also define marginal densities:

$$P_k(x_1, x_2, \ldots, x_k) \equiv \int\limits_K dx_{k+1} dx_{k+2} \cdots dx_N \ P_N(x_1, x_2, \ldots, x_N) \ , \tag{3}$$

for $k = 0, 1, 2, \ldots, N$, so that $P_0 = 1$. For truly random points, we have the ideal iid uniform distribution:

$$P_N(x_1, x_2, \ldots, x_N) = 1 \ . \tag{4}$$

We want, however, to be more general, and we shall write

$$P_k(x_1, x_2, \ldots, x_k) = 1 - \frac{1}{N} F_k(x_1, x_2, \ldots, x_k) \ . \tag{5}$$

Obviously, none of the F_k can exceed N, and for truly random points they are identically zero. Moreover, since the order in which the points enter is immaterial in this integration problem, we shall assume that all P_k and F_k are invariant under any permutation of the arguments.

We are now ready to estimate the error, by computing the various moments of the probability distribution of η. Denoting by brackets the average over P_N, we have for the first moment

$$\langle \eta \rangle \;=\; -J_1 + \frac{1}{N}\sum_{k=1}^{N}\langle f(x_k)\rangle \;=\; \frac{1}{N}\int\limits_K dx\; f(x)F_1(x) \;\; . \tag{6}$$

We see that, if the integration is to be unbiased for all integrands, we needs must have

$$F_1(x) = 0 \;\; , \tag{7}$$

which means that, *individually,* each integration point x_k in X_N must be uniformly distributed, and the difference between quasi-random and truly random point sets may show up only in correlations between points. Assuming this to be indeed the case, we then turn to the second moment:

$$\begin{aligned}\langle \eta^2 \rangle \;&=\; J_1^2 - \frac{2J_1}{N}\sum_{k=1}^{N}\langle f(x_k)\rangle + \frac{1}{N^2}\sum_{k,l=1}^{N}\langle f(x_k)f(x_l)\rangle \\ &=\; \frac{1}{N}\left(J_2 - J_1^2 - \left(1-\frac{1}{N}\right)\int\limits_K dx_1 dx_2\; f(x_1)f(x_2)F_2(x_1,x_2)\right) \;\; .\end{aligned} \tag{8}$$

The following facts become apparent from this result. In the first place, for truly random points, F_2 vanishes and we recover the standard Monte Carlo estimate. Secondly, only a small, $\mathcal{O}\,(1/N)$, deviation in P_N from the truly random uniform iid case can already significantly diminish the error, due to the delicate cancellations involved in the computation of $\langle \eta^2 \rangle$. Finally, the mechanics behind this improvement become obvious: in very uniform point sets (such as quasi-random point sets), that have a low discrepancy, the points are spread 'more evenly' than in typical random sets: they 'repel' each other, leading to a positive value of $F_2(x_1, x_2)$ whenever x_1 and x_2 are 'close' in K: this suppresses the contribution to $\langle \eta^2 \rangle$ from regions where $f(x)^2$ tends to be large. Conversely, for very non-uniform sets, where the points are more 'clumped' together, and whose discrepancy is large, F_2 will be negative for neighbouring points, with a corresponding punishment in the error estimate. A very simple and instructive illustration of this idea is given in Appendix B.

A final remark is in order here: for truly random point sets, the usefulness of the error estimate relies on the fact that the distribution of η tends to a Gaussian, as implied by the Central Limit Theorem. In principle, we ought to reconstruct a proof of this theorem for the more general P_N discussed here. Although we have not yet done so, we have good reason to believe that a

Central Limit Theorem holds in our case as well, with an implied Gaussian width following from Eq.(8).

2.2 Generating functions

Our task is now clear: we have to find, for quasi-random point sets, a workable definition of P_N, and a corresponding formula for F_2. To do so, we first assume that there exists *some* measure of (non-)uniformity of point sets X_N: that is, there must be given some *discrepancy* D_N as a function of the x_k, and we shall also assume that we can, for a given X_N, compute its value which we shall denote by s:

$$D_N(x_1, x_2, \ldots, x_N) = s \quad . \tag{9}$$

A 'small' value of s shall indicate a point set that is relatively uniform compared to a truly random one. We shall defer an explicit definition of D_N to the next section, and for the moment just assume that one is given. We then propose to use for P_N the probability density obtained by assuming that all points are uniformly distributed, with the additional constraint that the discrepancy takes on the value s:

$$\begin{aligned} P_N(s; x_1, \ldots, x_N) &\equiv \frac{H_N(s; x_1, \ldots, x_N)}{H_0(s)} \quad , \\ H_k(s; x_1, \ldots, x_k) &\equiv \int_K dx_{k+1} \cdots dx_N \; \delta(D_N(x_1, \ldots, x_N) - s) \quad , \end{aligned} \tag{10}$$

where we have introduced the Dirac delta distribution to handle the discrepancy constraint. The function $H_0(s)$ is, then, the probability distribution of s over the ensemble of all truly random point sets, which is an interesting quantity in its own right. The idea behind this definition is the following. We are not allowed, in principle, to consider a quasi-random point set as a 'typical' one in the whole random ensemble, since s is (hopefully) small compared to the expected value for random sets: but, in the subset with that particular value of s, it may very well be typical (indeed, this is the same fingers-crossed attitude that allows us to use, for pseudo-random point sets, the standard Monte Carlo error estimate in the first place). Moreover, if no information whatsoever is available on the value of s, we have to integrate over all s with probability density $H_0(s)$, upon which the delta function constraint drops out and we are back in the truly random case.

We now proceed to calculate F_2. Let us define a set of moment-generating functions as follows:

$$G_k(z; x_1, \ldots, x_k) \equiv \int\limits_K dx_{k+1} \cdots dx_N \ \exp\left(z D_N(x_1, \ldots, x_N)\right) \ , \tag{11}$$

that is, the generating function where the first k of the integration points are kept fixed, and the remaining ones are integrated over: $G_0(z)$, then, is the moment-generating function for $H_0(s)$. We employ the definition of the Dirac delta distribution as a Fourier/Laplace transform to write

$$H_k(s; x_1, \ldots, x_k) = \frac{1}{2\pi i} \int\limits_{-i\infty}^{i\infty} dz \ e^{-zs} G_k(z; x_1, \ldots, x_k) \ , \tag{12}$$

where the integration contour must run to the left of any singularities. So, knowledge of the G_k allows us to compute everything we need. We may uniquely split each G_k into a constant part, and a part that depends on the $x_{1,\ldots,k}$ and averages out to zero:

$$\begin{aligned} G_k(z; x_1, \ldots, x_k) &= G_0(z) + \frac{1}{N} \rho_k(z; x_1, \ldots, x_k) \ , \\ \int\limits_K dx_1 \cdots dx_k \ \rho_k(z; x_1, \ldots, x_k) &= 0 \ . \end{aligned} \tag{13}$$

It follows that we can express F_k as

$$\begin{aligned} F_k(s; x_1, \ldots, x_k) &= -R_k(s; x_1, \ldots, x_k)/H_0(s) \ , \\ R_k(s; x_1, \ldots, x_k) &= \frac{1}{2\pi i} \int\limits_{-i\infty}^{i\infty} dz \ e^{-zs} \rho_k(z; x_1, \ldots, x_k) \ . \end{aligned} \tag{14}$$

In this way, we can compute the necessary two-point correlation F_2, provided we can compute, at least in the form of an asymptotic expansion in $1/N$, the moment-generating functions G_k: the dominant contribution to F_k will then come from the leading terms in G_0 and ρ_k. This depends, of course, on a manageable choice of the discrepancy D_N, and this will be the subject of the next section. We wish to point out that, actually, an exact expression for the F_k would allow us to determine the lowest possible value for s, simply by putting $F_k(s; x_1, \ldots, x_k) = N$; from the fact that it is so hard to improve on the Roth bound, we may infer that an asymptotic expansion in $1/N$ is probably the best possible result we can hope for at this moment. Finally,

we remark that a check on our results for $R_2(s; x_1, x_2)$ is provided by the identities

$$\int_K dx_2 \; R_2(s; x_1, x_2) = 0 \quad , \quad \int_0^\infty ds \; R_2(s; x_1, x_2) = 0 \quad , \tag{15}$$

where the first identity follows from Eq.(7), and the second one by construction.

3 The Fourier discrepancy

We now turn to a useful definition of a discrepancy measure for X_N. This we do by first constructing a model of the class of integrands likely to be encountered, and then considering the squared integration error expected for this class of integrands. This procedure is analogous to that employed by Woźniakowski [10].

3.1 The one-dimensional case

For simplicity, we shall first discuss the case $D = 1$, and only further on generalize to more dimensions. We start by assuming that our integrands admit of a decomposition into a set of orthonormal functions. These satisfy

$$\int_K dx \; u_m(x) u_n(x) = \delta_{mn} \quad , \quad \sum_{n \geq 0} u_n(x) u_n(y) = \delta(x - y) \quad . \tag{16}$$

A general integrand $f(x)$ can then be written as

$$f(x) = \sum_{n \geq 0} v_n \; u_n(x) \quad , \tag{17}$$

so that the coefficients v_n determine the function. Since these coefficients form an enumerable set, it is fairly easy to set up a combined probability distribution for them, which gives us then a probability measure $\mathcal{D}f$ on the class of integrands. We take the following choice:

$$\mathcal{D}f = \prod_{n \geq 0} dv_n \; \frac{\exp(-v_n^2 / 2\sigma_n^2)}{\sqrt{2\pi\sigma_n^2}} \quad , \tag{18}$$

so that each coefficient v_n is distributed normally with mean value 0 and standard deviation σ_n. We shall call σ_n the *strength* of the mode $u_n(x)$ in

our function ensemble. Denoting by $\langle\rangle_f$ an average over the ensemble under the measure $\mathcal{D}f$, we then have

$$\langle v_n \rangle_f = 0 \quad , \quad \langle v_n v_m \rangle_f = \sigma_m^2 \delta_{mn} \quad . \tag{19}$$

We can also compute the moments of the integration error η over the integrand ensemble. The integral J_1 is of course just given by v_0, and moreover we have

$$\langle \eta \rangle_f = 0 \quad , \quad \langle \eta^2 \rangle_f = \frac{1}{N^2} \sum_{k,l=1}^{N} \beta_1(x_k, x_l) \quad , \tag{20}$$

where we have introduced functions β_m as

$$\beta_m(x,y) = \sum_{n>0} \sigma_n^{2m} u_n(x) u_n(y) \quad . \tag{21}$$

These functions, which will play an important rôle in our discussion, are symmetric in their arguments, and have the properties that

$$\int_K dy \; \beta_m(x,y) = 0 \quad , \quad \int_K dy \; \beta_m(x_1,y)\beta_n(x_2,y) = \beta_{m+n}(x_1,x_2) \quad , \tag{22}$$

for $m, n > 0$. In fact it is also simple to prove that the error η has a Gaussian distribution, not as a consequence of any law of large numbers, but because of the Gaussian character of our measure $\mathcal{D}f$ on the space of integrands.

Up to this point, we have not had to specify the particular orthonormal functions. In practice, we shall use the Fourier basis:

$$u_0(x) = 1 \quad , \quad u_{2n-1}(x) = \sqrt{2} \; \sin(2\pi n x) \quad , \quad u_{2n}(x) = \sqrt{2} \; \cos(2\pi n x) \; , \tag{23}$$

where n runs over the positive integers.

In addition, we shall make one very important assumption, namely that the sines and cosines have equal strength:

$$\sigma_{2n-1} \equiv \sigma_{2n} \quad . \tag{24}$$

We shall call this property *translational invariance*. Its physically reasonable motivation is the fact that under this assumption the mode with frequency n (made up from u_{2n-1} and u_{2n}) has a uniformly distributed phase. Translational invariance will enable us to considerably simplify our results. Most importantly, it leads us to write

$$\beta_m(x_1,x_2) = \sum_{n>0} 2(\sigma_{2n})^{2m} \cos(2\pi n(x_1 - x_2)) \quad , \tag{25}$$

so that the functions β only depend on the *difference* between x_1 and x_2, and we may also write $\beta_m(x_1 - x_2)$. Another attractive consequence is the fact that (as we shall see) point sets that only differ by a translation (assuming a periodic extension of K and X_N) have the same discrepancy, in contrast to, *e.g.*, the star discrepancy D_N^*.

The above consideration leads us to propose, as an appropriate definition of discrepancy, the following:

$$D_N = D_N(x_1, \ldots, x_N) \equiv \frac{1}{N} \sum_{k,l=1}^{N} \beta_1(x_k, x_l) \quad . \tag{26}$$

Note that we have taken out one factor of $1/N$: this will make the discrepancy independent of N for truly random points. It is easily seen that, for truly random points,

$$\langle D_N \rangle = \sum_{n>0} \sigma_n^2 \quad . \tag{27}$$

There are a few additional observations in order here. In the first place, note that the rôles of integrand and point set have, in some sense, been reversed here: in the derivation of the standard Monte Carlo error, we keep the integrand fixed, and average over the ensemble of point sets, upon which $\langle \eta^2 \rangle$ no longer depends on the particular point set but only upon a property of the integrand (namely its variance); while here we have kept the point set fixed, and averaged over an ensemble of integrands, so that $\langle \eta^2 \rangle_f$ no longer depends on the integrand, but only upon a property of the point set X_N, namely its discrepancy. A number of other results along these lines have been presented in [11]. In the second place, the zero mode with $n = 0$ does not enter in the discrepancy: this is reasonable since v_0 is in fact the desired integral, and the *error* just consists in our partial inability to get rid of the higher modes, with $n > 0$. Finally, the strengths σ_n cannot be chosen too arbitrarily. If our average integrand is to be quadratically integrable, we need to have

$$\left\langle \int_K dx \; f(x)^2 \right\rangle_f = \sum_{n \geq 0} \sigma_n^2 < \infty \quad , \tag{28}$$

while additional smoothness requirements will ask for an even faster convergence of this sum. An admissible, and reasonable, choice for the σ_n will be, for instance

$$\sigma_0 = 1 \;\; , \;\; \sigma_{2n} = 1/n \;\; , \;\; n > 0 \;\; , \tag{29}$$

and this we shall consider in the practical applications [9].

3.2 The more-dimensional case

The extension of the above discussion to more dimensions is quite straightforward. We shall use again x to denote D-dimensional points. The complete set of orthonormal functions is now enumerated not by a scalar n, but by a vector quantity $\vec{n}$ with indices $n^\mu \geq 0$, with $\mu = 1, 2, \ldots, D$. The set of orthonormal functions is now given by

$$u_{\vec{n}}(x) \equiv \prod_{\mu=1}^{D} u_{n^\mu}(x^\mu) \quad ; \tag{30}$$

the measure $\mathcal{D}f$ is straightforward, with the strengths of the various modes denoted by $\sigma_{\vec{n}}$, and the discrepancy is again given by

$$D_N = D_N(x_1, \ldots, x_N) \equiv \frac{1}{N} \sum_{k,l=1}^{N} \beta_1(x_k, x_l) \quad , \tag{31}$$

where

$$\beta_m(x, y) = \sum_{\vec{n}>0} \sigma_{\vec{n}}^{2m} u_{\vec{n}}(x) u_{\vec{n}}(y) \quad . \tag{32}$$

Here, the sum runs over all $\vec{n}$ except $\vec{n} = (0, 0, \ldots, 0)$. Again, for truly random points we have

$$\langle D_N \rangle = \sum_{\vec{n}>0} \sigma_{\vec{n}}^2 \quad . \tag{33}$$

Translational invariance in D dimensions requires

$$\begin{aligned} \sigma_{(2n^1, 2n^2, \ldots, 2n^D)} &= \sigma_{(2n^1-1, 2n^2, \ldots, 2n^D)} = \cdots \\ \cdots &= \sigma_{(2n^1, 2n^2, \ldots, 2n^D-1)} = \cdots \\ \cdots &= \sigma_{(2n^1-1, 2n^2-1, \ldots, 2n^D-1)} \quad , \end{aligned} \tag{34}$$

so that the strengths must be equal in groups of 2^{D-k}, where k is the number of vanishing components of $\vec{n}$. Again, for our integrands to be quadratically integrable on the average, we must have

$$\sum_{\vec{n}>0} \sigma_{\vec{n}}^2 < \infty \quad . \tag{35}$$

Now, we may choose to let the strengths be dominated by the highest partial frequency, for instance

$$\sigma_{\vec{n}} = \sigma(\max_\mu n^\mu) \quad . \tag{36}$$

Because of the increasing multiplicity with $\max_\mu n^\mu$, $\sigma_{\vec{n}}$ must then decrease more rapidly than in the one-dimensional case. Other reasonable choices,

such as $\sigma_{\vec{n}} = \sigma(\sum_\mu n^\mu)$, lead to the same conclusion. As usual, we encounter here the phenomenon that a smooth function in one dimension is not precisely the same concept as the projection onto one dimension of a function that is smooth in many dimensions: the 'curse of dimensionality' crops up again, in disguise.

We have now finished the first part of our program, that is, the establishing of a reasonable definition of a discrepancy. We want to remark that it differs from the better known *star-discrepancy* D_N^*, discussed so extensively in [12]. That one can, in fact, also be derived from a class of integrands, as first made explicit in [10]. The appropriate integrand class is defined by taking for $\mathcal{D}f$ the Wiener sheet measure; although mathematically attractive, this is by no means the preferred choice for many practical applications, since it singles out functions that are everywhere continuous but nowhere differentiable. Moreover, the translational invariance will lead, as we shall see, to $F_1(s; x_1) \equiv 0$, and the lack of this invariance for the star-discrepancy probably implies that we cannot simply prove that it leads to unbiased integral estimates. We should like to point out that, in fact, we have been able to derive the form of $H_0(s)$ for the star-discrepancy as well [13].

4 Discrepancies by Feynman diagrams

We must now proceed with our program, and find expressions for

$$G_k(z; x_1, \ldots, x_k) = \sum_{m \geq 0} \frac{z^m}{m!} \langle D_N^m \rangle_k \ ,$$

$$\langle D_N^m \rangle_k = \frac{1}{N^m} \int_K dx_{k+1} \cdots dx_N \sum_{r_{1,\ldots,2m}=1}^{N} \beta_1(x_{r_1}, x_{r_2}) \cdots \beta_1(x_{r_{2m-1}}, x_{r_{2m}}) \ . \quad (37)$$

We shall call the points $x_1, \ldots, x_k$, that are kept fixed, *external points*, and the remaining $N - k$ ones, that are integrated over, *internal points*. Note that, in the multiple sum, each index r_j runs over both external and internal points. The various internal points are essentially indistinguishable, and they give rise to combinatorial factors; the study of these factors will enable us to write $\langle D_N^m \rangle_k$ as an asymptotic series in $1/N$.

The part of the multiple sum where all indices $r_1, \ldots, r_{2m}$ are equal and internal gives rise to a single combinatorial factor $N-k$; if they are distributed over two different internal values, we shall have a combinatorial factor $(N-k)(N-k-1)$, and so on. It follows that the largest combinatorial factor that

can, in principle, occur, is the one for the case where all indices are internal and distinct. However, that contribution consists of separate factors

$$\int_K dx_{r_i} dx_{r_j} \beta_1(x_{r_i}, x_{r_j}) \ ,$$

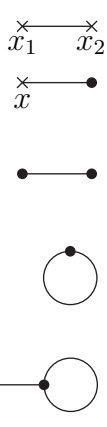
which vanish. Therefore, the *actually* largest combinatorial factor is that for the situation where the indices are all internal, and fall in distinct pairs: the corresponding combinatorial factor is

$$\begin{aligned} (N-k)^{\underline{m}} &= (N-k)!/(N-k-m)! \\ &= N^m - N^{m-1}\left[\frac{m^{\underline{2}}}{2} + km\right] + \mathcal{O}\left(N^{m-2}\right) \ . \end{aligned} \tag{38}$$

The 'falling power' notation is taken from Graham *et al.* [14]; its asymptotic expansion for large N is derived in Appendix A.

4.1 Feynman diagrams

We are now ready to compute the first few terms in the $1/N$ expansion of $\langle D_N^m \rangle_k$. We start by introducing a diagrammatical technique. Each internal point we shall denote by a dot, and each external point by a cross. Every function β_1 shall be denoted by a *link*, a solid line between dots or crosses. Every contribution to $\langle D_N^m \rangle_k$ will therefore contain precisely m links. Identical values of the summation indices mean that the corresponding points will be contracted, leading to *vertices* with two, three, four, ... legs. Neglecting for the moment the combinatorials we have, for instance,



$$\begin{aligned} \times\!\!-\!\!-\!\!-\!\!\times &= \beta_1(x_1, x_2) \ , \\ \times\!\!-\!\!-\!\!-\!\!\bullet &= \int_K dy \ \beta_1(x, y) = 0 \ , \\ \bullet\!\!-\!\!-\!\!-\!\!\bullet &= \int_K dy_1 dy_2 \ \beta_1(y_1, y_2) = 0 \ , \\ \bigcirc &= \int_K dy \ \beta_1(y, y) \ , \\ \bigcirc\!\!-\!\!\bullet\!\!-\!\!\bullet\!\!-\!\!\bigcirc &= \int_K dy_1 dy_2 \ \beta_1(y_1, y_1)\beta_1(y_1, y_2)\beta_1(y_2, y_2) \ , \end{aligned} \tag{39}$$

where the two zero results follow from Eq.(22). In each diagram or product of diagrams, the combinatorial factor can be read off immediately from the

number of internal points in evidence: if there are p internal points, the factor is $(N-k)^{\underline{p}}$. External points do not contribute a combinatorial factor, but it must be kept in mind that we shall have to sum over all external points. For instance, the diagram $\times\!\longrightarrow\!\times$ must be interpreted as

$$\times\!\longrightarrow\!\times \;=\; \sum_{\substack{k,l=1\\ k\neq l}}^{N} \; \underset{x_k}{\times}\!\longrightarrow\!\underset{x_l}{\times}\;, \tag{40}$$

so that this diagram actually contains $k(k-1)$ terms.

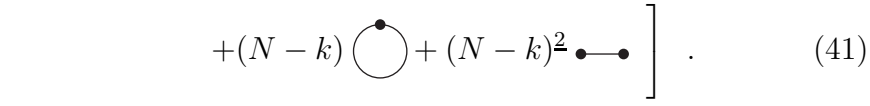
In writing out the $1/N$ expansion, we can simply determine which kinds of diagrams will contribute, as follows. The leading power of a two-point internal vertex is 1, that of a three-point internal vertex is $1/\sqrt{N}$, for a four-point internal vertex we have $1/N$, and so on. Formally, this is similar to a field theory with a universal coupling constant g proportional to $1/\sqrt{N}$. Moreover, each external point will effectively carry a factor $1/\sqrt{N}$. If we decide to keep only the terms up to and including $1/N$, we shall have to allow for at most two external points, or a single external point and one internal three-vertex, or two internal three-vertices, or a single internal four-vertex. In a field theory, this corresponds to the first-order correction, proportional to g^2. Moreover, our property of translational invariance implies precisely the analogue for a field theory, namely momentum conservation[2]. Also, the external points, that act somewhat like sources, need rescaling by factors $\sqrt{N}$, which could be envisioned as the truncation process by which a Green's function is turned into an S-matrix element. This is precisely what the extraction of the factor $1/N$ in Eq.(5) does for us, and so we recognize what the function $F_2(s;x_1,x_2)$ actually is: it is the full propagator.

4.2 An illustration: the first two moments

We shall now show how to apply the diagrammatic techniques in the calculation of $G_k(z;x_1,\ldots,x_k)$. The first nontrivial term comes from $\langle D_N\rangle_k$, which we can write diagrammatically as

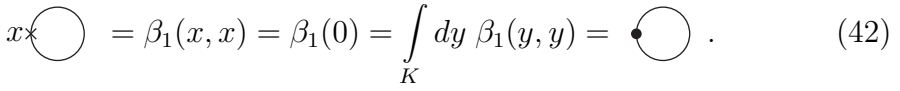
$$\langle D_N\rangle_k \;=\; \frac{1}{N}\Bigg[\; \times\!\longrightarrow\!\times \;+\; \bigcirc\!\!\!\!\times \;+\; 2(N-k)\,\times\!\longrightarrow\!\bullet$$

[2]The analogy with a real field theory cannot be carried too far, however, since such a theory has an infinite number of degrees of freedom. In the present case, all vertices and combinatorial factors also carry subleading contributions, which must be properly taken into account.



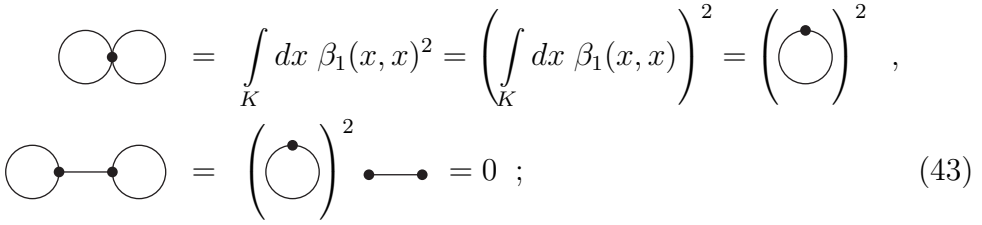
(41)

We may considerably simplify this. In the first place, we have $\times\!\!-\!\!\bullet = 0$ and $\bullet\!\!-\!\!\bullet = 0$, see Eq.(39). In addition, the property of translational invariance has the important consequence that a diagram with only a single external point evaluates to a constant. For instance,



(42)

This also implies that, whenever two parts of a diagram are connected by a single vertex, we may split it up (of course, without changing the combinatorial factor!), so that, for instance,



(43)

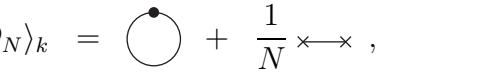
the last line is an example of the more general phenomenon that all *tadpole* diagrams with only internal points evaluate to zero. Again, this is due to our assumption of translational invariance: incidentally, it immediately proves that

$$G_1(z;x_1) = G_0(z) \quad \text{to all orders in } 1/N, \tag{44}$$

which in its turn implies that

$$F_1(s;x_1) = 0 \quad \text{to all orders in } 1/N, \tag{45}$$

as required in Eq.(7). We can now write $\langle D_N \rangle_k$ as



(46)

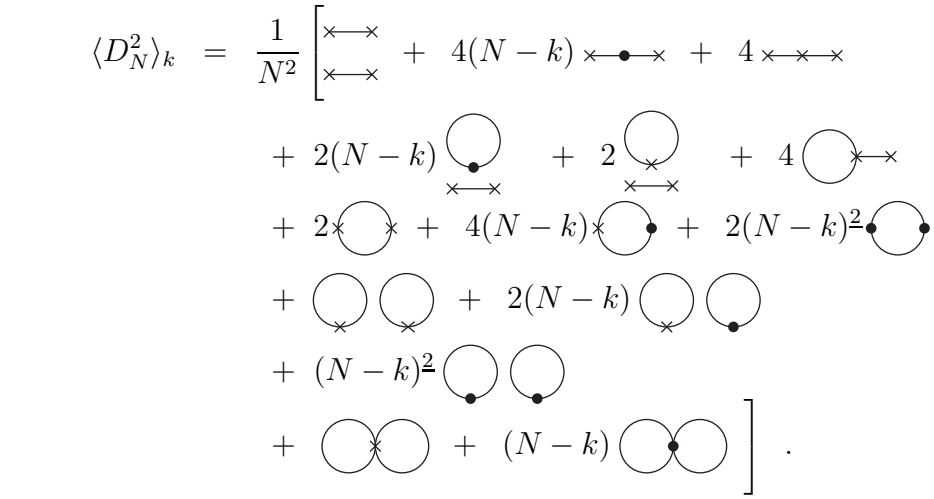
where we must keep in mind that the second term implies a summation over the $k(k-1)$ pairs of *different* points in $x_1, x_2, \ldots, x_k$. It follows trivially that, as required,

$$\int_K dx_k \, \langle D_N \rangle_k = \langle D_N \rangle_{k-1} \quad \Rightarrow$$

$$\Rightarrow \quad \int_K dx_k \, F_k(s;x_1,\ldots,x_k) = F_{k-1}(s;x_1,\ldots,x_{k-1}) \ . \tag{47}$$

We now proceed to the next order, and compute $\langle D_N^2 \rangle_k$. The only contributions that do not immediately vanish under Eq.(22) are

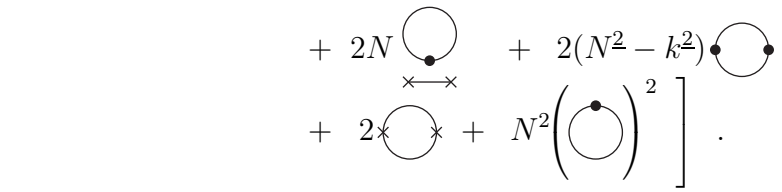
$$\langle D_N^2 \rangle_k = \frac{1}{N^2} \Bigg[$$



$$\Bigg] \quad . \tag{48}$$

Using translational invariance, we may rewrite this as

$$\langle D_N^2 \rangle_k = \frac{1}{N^2} \Bigg[$$



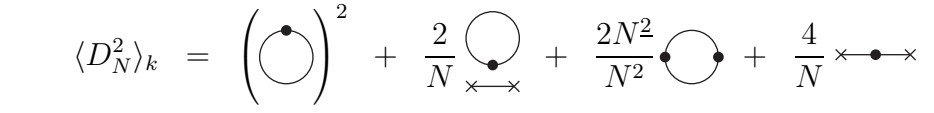
$$\Bigg] \quad . \tag{49}$$

It can again easily be checked that

$$\int_K dx_k \, \langle D_N^2 \rangle_k = \langle D_N^2 \rangle_{k-1} \quad . \tag{50}$$

If we restrict ourselves to terms of order $\mathcal{O}(1)$ and $\mathcal{O}(N^{-1})$, we have

$$\langle D_N^2 \rangle_k =$$



$$. \tag{51}$$

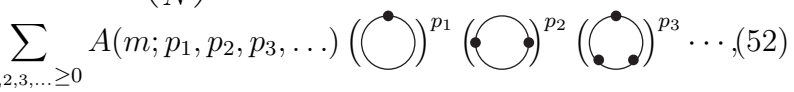
4.3 The general result to leading order

We shall now discuss the computation of all diagrams of the leading necessary order. To start, let us concentrate on the leading terms, of order $\mathcal{O}(1)$. These

are contributed by graphs containing only internal two-vertices, that is, they must be of the form of a product of closed loops:

$$G_k(z) = W_0(m) + \mathcal{O}\left(\frac{1}{N}\right) \ ,$$

$$W_0(m) \equiv \sum_{p_{1,2,3,\ldots}\geq 0} A(m;p_1,p_2,p_3,\ldots)$$



$$\cdots , \quad (52)$$

where we have left out the factor $N^{\underline{m}}/N^m$, and imply the constraint

$$m = p_1 + 2p_2 + 3p_3 + \cdots \ . \quad (53)$$

The factor $A(m;p_1,p_2,\ldots)$ is governed by a recursion relation. We may go from $m-1$ to m by either adding a single one-link loop, or by putting an extra link in any k-link loop, thereby turning it into a $(k+1)$-link loop. The recursion relation therefore reads

$$A(m;p_1,p_2,p_3,\ldots) = A(m-1;p_1-1,p_2,p_3,\ldots) \ + \\ \sum_{k>0} 2k(p_k+1)A(m-1;p_1,p_2,\ldots,p_k+1,p_{k+1}-1,\ldots) \ . \quad (54)$$

We may use the Ansatz

$$A(m;p_1,p_2,p_3,\ldots) = \frac{m!}{p_1!p_2!p_3!\cdots} C a_1^{p_1} a_2^{p_2} a_3^{p_3} \cdots \ , \quad (55)$$

with C and the a's to be determined. Putting this Ansatz in Eq.(54) leads us to

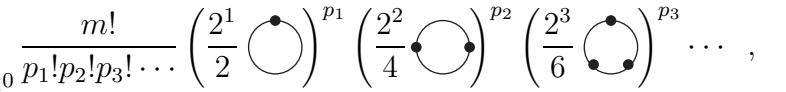
$$m = \frac{p_1}{a_1} + \sum_{k>0} 2k \frac{a_k}{a_{k+1}} p_{k+1} \ . \quad (56)$$

Since the coefficients of the p's are known from Eq.(53), this gives us the symmetry factor associated with each closed k-link loop:

$$a_k = \frac{2^k}{2k} \ . \quad (57)$$

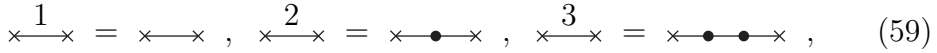
By inspection of $\langle D_N \rangle_k$, we also find $C = 1$. The result for the terms that give the leading contribution is therefore

$$W_0(m) \equiv \sum_{p_{1,2,3,\ldots}\geq 0} \frac{m!}{p_1!p_2!p_3!\cdots}$$



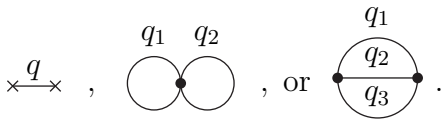
$$\cdots \ , \quad (58)$$

We now come to the subleading terms. Let us introduce the notation



$$, \quad (59)$$

so as to indicate the number of links between two vertices. The subleading contributions are characterised by the additional presence of one of the following diagrams:



It is actually only the first of these graphs that is needed for the purpose of computing ρ_k to first order: the other two will only contribute to subleading terms in $H_0(s)$ which are not relevant to the order we are working in here. The relevant contribution is, therefore

$$\rho_k(z; x_1, \ldots, x_k) = W_1(m) + \mathcal{O}\left(\frac{1}{N}\right) \quad ,$$

$$W_1(m) \equiv \sum_{\substack{p_{1,2,\ldots} \geq 0 \\ q>0}} B(m; q, p_1, p_2, \ldots) \overset{q}{\times\!\!\longrightarrow\!\!\times} \left(\bigcirc\right)^{p_1} \left(\bigcirc\right)^{p_2} \cdots \quad , \quad (60)$$

with

$$m = q + p_1 + 2p_2 + 3p_3 + \cdots \quad . \quad (61)$$

The coefficient B satisfies a recursion relation similar to that of A:

$$\begin{aligned} B(m; q; p_1, p_2, \ldots) &= 2qB(m-1; q-1, p_1, p_2, \ldots) \\ &+ B(m-1; q, p_1 - 1, p_2, \ldots) \\ &+ \sum_{k>0} 2k(p_k+1) B(m-1; q, p_1, p_2, \ldots, p_k + 1, p_{k+1} - 1, \ldots) \quad . \end{aligned} \quad (62)$$

By the same trick as above, we may solve this relation. The symmetry factors for the loop diagrams are of course the same, and the diagram $\overset{q}{\times\!\!\longrightarrow\!\!\times}$ has a symmetry factor 2^{q-1}. We therefore find

$$W_1(m) \equiv \sum_{\substack{p_{1,2,\ldots} \geq 0 \\ q>0}} \frac{m!}{p_1! p_2! \cdots} \left(\frac{2^q}{2} \overset{q}{\times\!\!\longrightarrow\!\!\times}\right) \left(\frac{2^1}{2} \bigcirc\right)^{p_1} \left(\frac{2^2}{4} \bigcirc\right)^{p_2} \cdots \quad , \quad (63)$$

In the above derivations, it may be noted that the various powers of 2 arise from the fact that each link has two endpoints, while the remaining symmetry factors are just those of diagrams with topologically indistinguishable points ($1/2n$ for an n-link loop, $1/2$ for the propagator diagram).

4.4 The generating functions

We can now write down immediately the leading form for G_0 and ρ_k, by taking the appropriate sums over m. To this leading order, we do not have

to worry about subleading terms in $(N-k)^{\underline{m}}$ and $(N-k)^{\underline{m-1}}$. We have

$$G_0(z) \sim \sum_{m\geq 0} \frac{z^m}{m!}\langle D_N^m\rangle_0 = \sum_{m\geq 0}\frac{z^m}{m!}W_0(m)$$

$$= \exp\left(\frac{2z}{2}\bigcirc + \frac{(2z)^2}{4}\bigcirc\!\!\bigcirc + \frac{(2z)^3}{6}\bigcirc\!\!\bigcirc\!\!\bigcirc + \cdots\right) \quad , \quad (64)$$

and

$$\rho_k(z;x_1,\ldots,x_k) \sim \sum_{m\geq 1}\frac{z^m}{m!}W_1(m) = \frac{1}{2}\sum_{q>0}\left((2z)^q \overset{q}{\times\!\!\longrightarrow\!\!\times}\right) G_0(z) \ . \quad (65)$$

To evaluate the various diagrams we employ the functions β. From the following propagator diagrams, with endpoints x and y:

$$\begin{aligned}
\overset{1}{\times\!\!\longrightarrow\!\!\times} &= \beta_1(x,y) \ , \\
\overset{2}{\times\!\!\longrightarrow\!\!\times} &= \int_K dz\ \beta_1(x,z)\beta_1(z,y) = \beta_2(x,y) \ , \\
\overset{3}{\times\!\!\longrightarrow\!\!\times} &= \int_K dz\ \beta_2(x,z)\beta_1(z,y) = \beta_3(x,y) \ , \\
&\vdots \\
\overset{q}{\times\!\!\longrightarrow\!\!\times} &= \beta_q(x,y) = \sum_{\vec{n}>0}(\sigma_{\vec{n}})^{2q}\ u_{\vec{n}}(x)u_{\vec{n}}(y) \ ,
\end{aligned} \quad (66)$$

we find the following representations:

$$\phi(z;x,y) \equiv \sum_{q\geq 1}(2z)^q \overset{q}{\times\!\!\longrightarrow\!\!\times} = \sum_{\vec{n}>0}\frac{2z\sigma_{\vec{n}}^2}{1-2z\sigma_{\vec{n}}^2}\ u_{\vec{n}}(x)u_{\vec{n}}(y) \ , \quad (67)$$

$$\log G_0(z) = -\frac{1}{2}\sum_{\vec{n}>0}\log\left(1-2z\sigma_{\vec{n}}^2\right) \ , \quad (68)$$

and

$$\rho_k(z;x_1,\ldots,x_k) = \frac{1}{2}\sum_{i\neq j}\phi(z;x_i,x_j)G_0(z) \ , \quad (69)$$

with the indices i and j running from 1 to k. For the case $k=2$ in which we are interested, this specializes to

$$\rho_2(z;x_1,x_2) = \phi(z;x_1-x_2)G_0(z) \ . \quad (70)$$

In the language of field theory, the function $\rho_2(z;,x_1,x_2)/N$ may be recognized as the Dyson-summed, or *dressed*, two-point Green's function, including vacuum diagrams $G_0(z)$. Here $G_0(z)$ plays the rôle of the path integral in the free-field approximation.

5 Conclusions

In this first paper we have addressed the question of how, given information on the uniformity property of a particular point set X_N in terms of a discrepancy (defined in Eq.(31)), we may hope to improve on the error estimate when we apply the point set X_N in numerical integration. To this end, we need the function $F_2(s; x_1, x_2)$ which is defined as the ratio of two functions, $R_2(s; x_1, x_2)$ and $H_0(s)$, which are themselves again given as the Fourier transforms of $\rho_k(z; x_1, x_2)$ and $G_0(z)$, respectively. We have developed a diagrammatic approach that allows us to systematically compute a series expansion for these two objects in powers of $1/N$. We have derived explicit expressions for the leading term in these expansions: they are given by Eq.(69) and Eq.(68), respectively.

Of course, the above discussion has been purely formal. We have been able to point out interesting parallels between our results and properties of a classical field theory; but all this will be moot unless we can, in fact, turn what we have learned into an expression for $F_2(s; x_1, x_2)$ that is explicit and simple to evaluate. This will be the subject of further publications in this series.

Apart from the explicit expressions for $\langle D_N \rangle_k$ and $\langle D_N^2 \rangle_k$, our results have been strictly limited to the leading terms, that are independent of N. Since any numerical integration usually employs a large value for N anyway, we feel that we are justified in this. However, there are issues for which a further expansion ought to be useful. For instance, it would be interesting to see how far we could approximate a lower bound on D_N à la Roth, by either computing further terms in F_2 and putting its value to N, or by trying to establish bounds on that value for s where $H_0(s)$ vanishes, for finite N. These issues are beyond our scope at this moment, but we feel that we have at least laid some of the groundwork in this paper.

Appendix A: Expansion of falling powers

In this appendix we give a simple derivation of the asymptotic expansion of the quantity $(N-k)^{\underline{m}}$. It is based upon the binomial-theorem fact

$$\sum_{m \geq 0} \frac{1}{m!} a^{\underline{m}} x^m = (1+x)^a \quad . \tag{71}$$

We therefore have

$$\sum_{m \geq 0} \frac{1}{m!} (N-k)^{\underline{m}} \left(\frac{x}{N}\right)^m = \left(1 + \frac{x}{N}\right)^{(N-k)} \quad . \tag{72}$$

Now, one the one hand, we have

$$\left(1+\frac{x}{N}\right)^N = e^x\left(1-\frac{x^2}{2N}+\frac{x^3}{3N^2}+\frac{x^4}{8N^2}+\mathcal{O}\left(N^{-3}\right)\right) \quad , \tag{73}$$

and, on the other hand,

$$\left(1+\frac{x}{N}\right)^{-k} = 1-\frac{kx}{N}+\frac{k(k+1)x^2}{2N^2}+\mathcal{O}\left(N^{-3}\right) \quad . \tag{74}$$

Multiplying the expansions (74) and (73), and carefully keeping track of the coefficient of x^m, we then find

$$\begin{aligned}\frac{(N-k)^{\underline{m}}}{N^m} &= 1-\frac{1}{N}\left[\frac{1}{2}m^{\underline{2}}+km\right] \\ &\quad +\frac{1}{N^2}\left[\frac{1}{8}m^{\underline{4}}+\frac{1}{3}m^{\underline{3}}+\frac{1}{2}km^{\underline{3}}+\frac{1}{2}k(k+1)m^{\underline{2}}\right] \\ &\quad +\mathcal{O}\left(N^{-3}\right) \quad ,\end{aligned} \tag{75}$$

and, of course, also the subleading expressions

$$\begin{aligned}(N-k)^{\underline{m-1}}/N^m &= \frac{1}{N}-\frac{1}{N^2}\left[\frac{1}{2}m^{\underline{2}}+(k-1)(m-1)\right]+\mathcal{O}\left(N^{-3}\right) \quad , \\ (N-k)^{\underline{m-2}}/N^m &= \frac{1}{N^2}+\mathcal{O}\left(N^{-3}\right) \quad .\end{aligned} \tag{76}$$

Obviously, this expansion can be continued *ad nauseam*; moreover, it also provides a way of computing complicated constraint sums of products like

$$\sum_{k_1,k_2,\ldots,k_p=1}^{N-k} k_1 k_2 \cdots k_p \quad ,$$

with the constraints that all indices must be different.

Appendix B: an illustrative model

In this appendix we show how error improvement arises in a very simple one-dimensional model. We assume that the functions to be integrated have only modes up to n, and that those modes have equal strength. That is, we take

$$\begin{aligned}\sigma_k^2 = 1/2 \quad &, \quad k = 1,2,\ldots,2n \quad , \\ \sigma_k^2 = 0 \quad &, \quad k > 2n \quad .\end{aligned} \tag{77}$$

For this extremely simple case, we have

$$G_0(z) = \frac{1}{(1-z)^n} \quad , \quad \phi(z;x) = \frac{2z}{1-z}\sum_{k=1}^{n}\cos(2\pi kx) \ . \tag{78}$$

We may easily compute the Laplace/Fourier transforms, and find

$$H_0(s) = \frac{1}{(n-1)!}s^{n-1}e^{-s} \quad , \quad F_2(s;x) = 2\left(1-\frac{s}{n}\right)\sum_{k=1}^{n}\cos(2\pi kx) \ . \tag{79}$$

The average discrepancy is $\langle s\rangle = n$ for truly random points. A typical integrand is, in this model, of the form

$$f(x) = \sum_{k=1}^{n} v_k \ \cos(2\pi k(x+\alpha_k)) \ , \tag{80}$$

with arbitrary v_k and α_k. We immediately find

$$J_1 = 0 \quad , \quad J_2 = \frac{1}{2}\sum_{k=1}^{n} v_k^2 \ , \tag{81}$$

and

$$\int\limits_K dx_1 dx_2 \ f(x_1)f(x_2)F_2(s;x_1-x_2) = \left(1-\frac{s}{n}\right)\frac{1}{2}\sum_{k=1}^{n} v_k^2 \ . \tag{82}$$

We conclude that, with the inclusion of the two-point correlation function F_2, the straightforward Monte Carlo error estimate is changed into the Quasi-Monte Carlo one, as follows:

$$\begin{aligned} &\langle\eta^2\rangle_{\text{Monte Carlo}} \ \rightarrow \\ &\rightarrow \ \langle\eta^2\rangle_{\text{Quasi-Monte Carlo}} \ = \ \frac{s}{\langle s\rangle}\langle\eta^2\rangle_{\text{Monte Carlo}} \ + \ \mathcal{O}\left(\frac{1}{N^2}\right) \ . \end{aligned} \tag{83}$$

We see that the error is improved if the actual discrepancy is small compared to its expected value.

References

[1] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods* (SIAM, Philadelphia, 1992) ch. 5., and references quoted therein.

[2] J.F. Koksma, *Mathematica B* (Zutphen), **11** (1942/1943) 7;
E. Hlawka, *Ann. Math. Pura Appl.* **54** (1961) 325;
H. Niederreiter, in: *Applications of Number Theory to Numerical Analysis*, K. Zaremba, ed. (Academic Press, New York, 1972) 203;
P.D. Proinov, *J. Approx. Theory* **52** (1988) 121.

[3] K.F. Roth, *Mathematika* **1** (1954) 73.

[4] J.G. van der Corput, *Nederl. Akad. Wetensch. Proc. ser. B* **38** (1935) 813, 1058;
J.H. Halton, *Numer. Math.* **2** (1960) 84; Err. 196.

[5] H. Faure, *Acta Arithmetica* **41** (1982) 337;
B.L. Fox, *ACM Trans. Math. Software* **12** (1986) 362.

[6] I.M. Sobol', *USSR Comp. and Math. Phys.* **7** (1967) 86;
P. Bratley and B.L. Fox, *ACM Trans. Math. Software* **14** (1988) 88.

[7] H. Niederreiter, *op. cit.*, ch. 4.;

[8] M. Berblinger and C. Sschlier, *Comp. Phys. Comm.* **66** (1991) 157.

[9] J.K. Hoogland and R. Kleiss, in preparation.

[10] H. Woźniakowski, *Bull. AMS* **24** (1991) 185.

[11] R. Kleiss, *Comp. Phys. Comm.* **71** (1992) 39.

[12] H. Niederreiter, *op. cit.*, ch. 2;
L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences* (Wiley, New York, 1974).

[13] F. James, J.K. Hoogland and R. Kleiss, in preparation.

[14] R.L. Graham, D.E. Knuth and O. Patashnik, *Concrete Mathematics* (Addison-Wesley, 1989).